PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/29, 15/82, C07K 14/415, G01N 33/53, A01N 65/00, A01H 5/00

A2

(11) International Publication Number: WO 99/13083

(43) International Publication Date: 18 March 1999 (18.03.99)

(21) International Application Number: PCT/EP98/05673

(22) International Filing Date: 7 September 1998 (07.09.98)

(30) Priority Data:
97202742.9 5 September 1997 (05.09.97) EP

(71) Applicant *(for all designated States except US)*: CROPDESIGN N.V. [BE/BE]; Technologiepark Zwijnaarde 3, B–9052 Gent (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: INZÉ, Dirk [BE/BE]; Driesstraat 18, B–9310 Moorsel–Aalst (BE). SEGERS, Gerda [BE/BE]; Lentestraat 21, B–9000 Gent (BE). DE VEYLDER, Lieven [BE/BE]; Popperodedries 42, B–9300 Aalst (BE).

(74) Agent: VOSSIUS & PARTNER; Postfach 86 07 67, D–81634 München (DE).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: METHOD AND MEANS FOR MODULATING PLANT CELL CYCLE PROTEINS AND THEIR USE IN CONTROLLING PLANT CELL GROWTH

(57) Abstract

Provided are DNA sequences encoding cell cycle interacting proteins as well as to methods for obtaining the same. Furthermore, vectors comprising said DNA sequences are described, wherein the DNA sequences are operatively linked to regulatory elements allowing expression in prokaryotic and/or eukaryotic host cells. In addition, proteins encoded by said DNA sequences, antibodies to said proteins and methods for their production are provided. Also described is a method for controlling or altering growth characteristics of a plant and/or a plant cell comprising introduction and/or expression of one or more cell cycle regulatory proteins functional in a plant or parts thereof and/or one or more DNA sequences encoding such proteins. Also provided is a process for disruption plant cell division by interfering in the expression of a substrate for cyclin–dependent protein kinase using a DNA sequence according to the invention wherein said plant cell is part of a transgenic plant. Further described are diagnostic compositions comprising the aforementioned DNA sequences, proteins and antibodies. Methods for the identification of compounds being capable of activating or inhibiting the cell cycle are described as well. Furthermore, transgenic plant cells, plant tissue and plants containing the above–described DNA sequences and vectors are described as well as the use of the aforementioned DNA sequences, vectors, proteins, antibodies and/or compounds identified by the method of the invention in plant cell and tissue culture, plant breeding and/or agriculture.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

Method and means for modulating plant cell cycle proteins and their use in controlling plant cell growth

The present invention relates to DNA sequences encoding cell cycle interacting proteins as well as to methods for obtaining the same. The present invention also provides vectors comprising said DNA sequences, wherein the DNA sequences are operatively linked to regulatory elements allowing expression in prokaryotic and/or eukaryotic host cells. In addition, the present invention relates to the proteins encoded by said DNA sequences, antibodies to said proteins and methods for their production. The present invention also relates to a method for controlling or altering growth characteristics of a plant and/or a plant cell comprising introduction and/or expression of one or more cell cycle regulatory proteins functional in a plant or parts thereof and/or one or more DNA sequences encoding such proteins. Also provided by the present invention is a process for disruption plant cell division by interfering in the expression of a substrate for cyclin-dependent protein kinase using a DNA sequence according to the invention wherein said plant cell is part of a transgenic plant. The present invention further relates to diagnostic compositions comprising the aforementioned DNA sequences, proteins and antibodies. The present invention also relates to methods for the identification of compounds being capable of activating or inhibiting the cell cycle. Furthermore, the present invention relates to transgenic plant cells, plant tissue and plants containing the above-described DNA sequences and vectors as well as to the use of the aforementioned DNA sequences, vectors, proteins, antibodies and/or compounds identified by the method of the invention in plant cell and tissue culture, plant breeding and/or agriculture.

Cell division is fundamental for growth in humans, animals and plants. Prior to dividing in two daughter cells, the mother cell needs to replicate its DNA. The cell cycle is traditionally divided into 4 distinct phases:

G1: the gap between mitosis and the onset of DNA synthesis;

S : the phase of DNA synthesis;

G2: the gap between S and mitosis;

M : mitosis, the process of nuclear division leading up to the actual cell division.

The distinction of these 4 phases provides a convenient way of dividing the interval between successive divisions. Although they have served a useful purpose, a recent flurry of experimental results, much of it as a consequence of cancer research, has resulted in a more intricate picture of the cell cycle's "four seasons" (Nasmyth, Science 274, 1643-1645, 1996; Nurse, Nature, 344, 503-508, 1990). The underlying mechanism controlling the cell cycle control system has only recently been studied in greater detail. In all eukaryotic systems, including plants, this control mechanism is based on two key families of proteins which regulate the essential process of cell division, namely protein kinases (cyclin dependent kinases or CDKs) and their activating associated subunits, called cyclins. The activity of these protein complexes is switched on and off at specific points of the cell cycle. Particular CDK-cyclin complexes activated at the G1/S transition trigger the start of DNA replication. Different CDK-cyclin complexes are activated at the G2/M transition and induce mitosis leading to cell division. Each of the CDK-cyclin complexes execute their regulatory role via modulating different sets of multiple target proteins. Furthermore, the large variety of developmental and environmental signals affecting cell division all converge on the regulation of CDK activity. CDKs can therefore be seen as the central engine driving cell division.

In animal systems and in yeast, knowledge about cell cycle regulations is now quite advanced. The activity of CDK-cyclin complexes is regulated at five levels: (i) transcription of the CDK and cyclin genes; (ii) association of specific CDK's with their specific cyclin partner; (iii) phosphorylation/dephosphorylation of the CDK and cyclins; (iv) interaction with other regulatory proteins such as SUC1/CKS1 homologues and cell cycle kinase inhibitors (CKI); and (v) cell cycle phase-dependent destruction of the cyclins and CKIs.

The study of cell cycle regulation in plants has lagged behind that in animals and yeast. Some basic mechanisms of cell cycle control appear to be conserved among eukaryotes, including plants. Plants were shown to also possess CDK's, cyclins and CKI's. However plants have unique developmental features which are reflected in

specific characteristics of the cell cycle control. These include for instance the absence of cell migration, the formation of organs throughout the entire lifespan from specialized regions called *meristems*, the formation of a cell wall and the capacity of non-dividing cells to re-enter the cell cycle. Another specific feature is that many plant cells, in particular those involved in storage (e.g. endosperm), are polyploid due to rounds of DNA synthesis without mitosis. This so-called endoreduplication is intimately related with cell cycle control.

Due to these fundamental differences, multiple components of the cell cycle of plants are unique compared to their yeast and animal counterparts. For example, plants contain a unique class of CDKs, such as CDC2b in *Arabidopsis*, which are both structurally and functionally different from animal and yeast CDKs.

The further elucidation of cell cycle regulation in plants and its differences and similarities with other eukaryotic systems is a major research challenge. Strictly for the case of comparison, some key elements about yeast and animal systems are described below in more detail.

As already mentioned above, the control of cell cycle progression in eukaryotes is mainly exerted at two transition points: one in late G_1, before DNA synthesis, and one at the G_2/M boundary. Progression through these control points is mediated by cyclin-dependent protein kinase (CDK) complexes, which contain, in more detail, a catalytic subunit of approximately 34-kDa encoded by the *CDK* genes. Both *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* only utilize one *CDK* gene for the regulation of their cell cycle. The kinase activity of their gene products $p34^{CDC2}$ and $p34^{CDC28}$ in *Sch. pombe* and in *S. cerevisiae*, respectively, is dependent on regulatory proteins, called cyclins. Progression through the different cell cycle phases is achieved by the sequential association of $p34^{CDC2/CDC28}$ with different cyclins. Although in higher eukaryotes this regulation mechanism is conserved, the situation is more complex since they have evolved to use multiple CDKs to regulate the different stages of the cell cycle. In mammals, seven CDKs have been described, defined as CDK1 to CDK7, each binding a specific subset of cyclins.

In animal systems, CDK activity is not only regulated by its association with cyclins but also involves both stimulatory and inhibitory phosphorylations. Kinase activity is positively regulated by phosphorylation of a Thr residue located between amino

acids 160-170 (depending on the CDK protein). This phosphorylation is mediated by the CDK-activating kinase (CAK) which interestingly is a CDK/cyclin complex itself. Inhibitory phosphorylations occur at the ATP-binding site (the Tyr15 residue together with Thr14 in higher eukaryotes) and are carried out by at least two protein kinases. A specific phosphatase, CDC25, dephosphorylates these residues at the G_2/M checkpoint, thus activating CDK activity and resulting in the onset of mitosis. CDK activity is furthermore negatively regulated by a family of mainly low-molecular weight proteins, called cyclin-dependent kinase inhibitors (CKIs). Kinase activity is inhibited by the tight association of these CKIs with the CDK/cyclin complexes.

With respect to cell cycle regulation in plants a summary of the state of the art is given below. In *Arabidopsis*, thusfar only two *CDK* genes have been isolated, *CDC2aAt* and *CDC2bAt*, of which the gene products share 56% amino acid identity. Both CDKs are distinguished by several features. First, only *CDC2aAt* is able to complement yeast $p34^{CDC2/CDC28}$ mutants. Second, CDC2aAt and CDC2bAt bear different cyclin-binding motifs (PSTAIRE and PPTALRE, respectively), suggesting they may bind distinct types of cyclins. Third, although both *CDC2aAt* and *CDC2bAt* show the same spatial expression pattern, they exhibit a different cell cycle phase-specific regulation. The *CDC2aAt* gene is expressed constitutively throughout the whole cell cycle. In contrast, *CDC2bAt* mRNA levels oscillate, being most abundant during the S and G_2 phases. In addition, multiple cyclins have been isolated from *Arabidopsis*. The majority displays the strongest sequence similarity with the animal A- or B-type class of cyclins, but also D-type cyclins have been identified. Although the classification of *Arabidopsis* cyclins is mainly based upon sequence similarity, limited data suggests that this organization corresponds with differential functions of each cyclin class. Direct binding of any cyclin with an *Arabidopsis* CDK subunit has, however, not yet been demonstrated.

In order to manage problems related to plant growth, plant architecture and/or plant diseases, it is believed to be of utmost importance to identify and isolate plant genes and gene products involved in the regulation of the plant cell division, and more particularly coding for and interacting with CDK's and/or their interacting proteins, responsible for the control of the cell cycle and the completion of the S and M phase of the cell cycle. If such novel genes and/or proteins have been isolated and

analyzed, the growth of the plant as a whole can be influenced. Also, the growth of specific tissues or organs and thus the architecture of the plant can be modified.

Thus, the technical problem underlying the present invention is to provide means and methods for modulating cell cycle proteins that are particular useful in agriculture and plant cell and tissue culture.

The solution to the technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the invention relates to a DNA sequence encoding a cell cycle interacting protein or encoding an immunologically active and/or functional fragment of such a protein, selected from the group consisting of:

(a) DNA sequences comprising a nucleotide sequence encoding a protein comprising the amino acid sequence as given in SEQ ID NO: 2;

(b) DNA sequences comprising a nucleotide sequence as given in SEQ ID NO: 1;

(c) DNA sequences hybridizing with the complementary strand of a DNA sequence as defined in (a) or (b) and encoding an amino acid sequence which is at least 80% identical to the amino acid sequence encoded by the DNA sequence of (a) or (b);

(d) DNA sequences, the nucleotide sequence of which is degenerated as a result of the genetic code to a nucleotide sequence of a DNA sequence as defined in any one of (a) to (c); and

(e) DNA sequences encoding a fragment of a protein encoded by a DNA sequence of any one of (a) to (d).

The term "cell cycle interacting protein" as denoted herein means a protein capable of binding to cyclin dependent kinases, in particular plant cyclin dependent kinases.

The term "cell cycle" means the cyclic biochemical and structural events associated with growth of cells, and in particular with the regulation of the replication of DNA and mitosis. The cycle is divided into periods called: G_0, Gap_1 (G_1), DNA synthesis (S), Gap_2 (G_2), and mitosis (M).

The terms "gene(s)", "polynucleotide", "nucleic acid sequence", "nucleotide sequence", "DNA sequence" or "nucleic acid molecule(s)" as used herein refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA, and RNA. It also includes known types of modifications, for example, methylation, "caps" substitution of one or more of the naturally occuring nucleotides with an analog. Preferably, the DNA sequence of the invention comprises a coding sequence encoding the above defined cell cycle interacting protein.

A "coding sequence" is a nucleotide sequence which is transcribed into mRNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to mRNA, cDNA, recombinant nucleotide sequences or genomic DNA, while introns may be present as well under certain circumstances.

Studies with a two-hybrid system which had been carried out in accordance with the present invention a new gene product interacting with CDC2aAt indicative of a hitherto unknown plant cell cycle regulatory nucleotide sequence was identified (further also called clone th65). The coding nucleotide sequence (reading frame) for the isolated clone th65 in SEQ.ID.NO.1 starts at position 3 and terminates at codon AGG (position 1184). The N-terminus contained a region rich in glutamine residues (SEQ ID NO 2.). Gln-rich domains are often part of the transcriptional activation domain of DNA binding factors (Mitchell and Tjian, 1989, Science, 245, 371-378) and have also been shown to be involved in protein-protein interactions (Bao *et al.*, 1996, PNAS, 93, 5037-5042). The th65 open reading frame also contains three consensus CDK phosphorylation sites. The identification of th65 as a CDC2aAt-associated protein and the presence of these phosphorylation sites indicates that the th65 protein is a substrate for CDKs.

Using a nucleic acid amplification technology, such as the polymerase chain reaction (PCR), a genomic DNA fragment can be isolated comprising the sequence

defined in SEQ.ID.NO.1. Thus a novel plant nucleotide sequence and polypeptide sequence, having a molecular weight of about 44 kDa, are provided. A homology search in databases showed a significant homology to a plant kinesin-related motor protein. The homology search was performed with the program BLASTN (version 2.0a19MP-WashU [build decunix3.2 01:53:29 05-feb-1998] (see Altschul, Nucleic Acids Res. 25 (1997), 3389-3402) on the Arabidopsis thaliana nucleic acids database at ATDB at Stanford (http://genome-www2.stanford.edu/cgi-bin/AtDB/nph-blast2atdb). The mRNA of th65 (AJ001729) was submitted to BLASTN and revealed homology to the genomic sequence from Arabidopsis thaliana (AB011479) (P(N)=4.4e-118). The function COMPARE (from the GCG 9.1 package, Genetics Computer Group Inc., Madison, USA) has been used to quantify the percentage of homology and similarity. With the parameters Gap weight = 12 and Length weight = 2 the function COMPARE resulted in an alignment showing 84.638 % similarity and 79.420 % identity. The genomic sequence (AB011479, clone MNA5) has been retrieved from the KAOS server with its annotations. A kinesin-like protein c was predicted on that sequence (73733..80900) as gene MNA5.12, having 22 exons and no homologue EST. The protein sequence (as given on the KAOS server) was then used to perform a BLASTP (version 2.0.4 [feb-24-1998]) with BEAUTY post-processing provided by the Human Genome Center, Baylor College of Medicine against the National Center for Biotechnology Information's non-redundant protein database (http://dot.imgen.bcm.tmc.edu:9331/seq-search/protein-search.html) and revealed homology to numerous kinesins from Arabidopsis and other organisms.

Kinesins and kinesin-related proteins are microtubule motor proteins involved in vesicle transport, spindle assembly and chromosome segregation at meiosis and mitosis. The present demonstration that CDC2a interacts with a putative kinesin-related motor proteins and the presence of consensus sites for CDK phosphorylation in the identified th65 clone indicates that CDKs directly modify the cytoskelet through phosphorylation of kinesin-related motor proteins. Thus it is expected that the nucleic acid molecules of the invention encode proteins that beside their intrinsic capability of interacting with cell cycle proteins in addition display the biological activity of kinesin-related motor proteins.

With "kinesin" is meant the superfamily of microtubule-based motor proteins which includes both plus- and minus-end-directed varieties and is widely distributed in microtubule-containing cells. Functions of kinesins may include membrane-bound organelle movement and mitosis. Also used specifically for the defining member of the superfamily (other members are considered to be kinesin-related proteins).

With "motor proteins" is meant mechanochemical enzymes involved in locomotion or transport.

With "mechanochemical enzyme" is meant an enzyme that converts chemical energy in the form of nucleoside triphosphates to mechanical energy such as force or mortility.

The present invention also relates to nucleic acid molecules hybridizing with the above-described nucleic acid molecules and differ in one or more positions in comparison with these as long as they encode a cell cycle interacting protein. By "hybridizing" it is meant that such nucleic acid molecules hybridize under conventional hybridization conditions, preferably under stringent conditions such as described by, e.g., Sambrook (Molecular Cloning; A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). An example of one such stringent hybridization condition is hybridization at 4XSSC at 65 °C, followed by a washing in 0.1XSSC at 65 °C for one hour. Alternatively, an exemplary stringent hybridization condition is in 50 % formamide, 4XSSC at 42 °C. Cell cycle interacting proteins derived from other organisms such as mammals, in particular humans, may be encoded by other DNA sequences which hybridize to the sequences for plant cell cycle interacting proteins under relaxed hybridization conditions and which code on expression for peptides having the ability to interact with cell cycle proteins. Examples of such non-stringent hybridization conditions are 4XSSC at 50 °C or hybridization with 30-40 % formamide at 42 °C. Such molecules comprise those which are fragments, analogues or derivatives of the cell cycle interacting protein of the invention and differ, for example, by way of amino acid and/or nucleotide deletion(s), insertion(s), substitution(s), addition(s) and/or recombination(s) or any other modification(s) known in the art either alone or in combination from the above-described amino acid sequences or their underlying

nucleotide sequence(s). Methods for introducing such modifications in the nucleic acid molecules according to the invention are well-known to the person skilled in the art. The invention also relates to nucleic acid molecules the sequence of which differs from the nucleotide sequence of any of the above-described nucleic acid molecules due to the degeneracy of the genetic code. All such fragments, analogues and derivatives of the protein of the invention are included within the scope of the present invention, as long as the essential characteristic immunological and/or biological properties as defined above remain unaffected in kind, that is the novel nucleic acid molecules of the invention include all nucleotide sequences encoding proteins or peptides which have at least a part of the primary structural conformation for one or more epitopes capable of reacting with antibodies to cell cycle interacting proteins which are encodable by a nucleic acid molecule as set forth above and which have comparable or identical characteristics in terms of binding to cyclin dependent kinases, in particular plant cyclin dependent kinases. Part of the invention is therefore also nucleic acid molecules encoding a polypeptide comprising at least a functional part of a cell cycle interacting protein encoded by a nucleic acid sequence comprised in a nucleic acid molecule according to the invention. An example for this is that the polypeptide or a fragment thereof according to the invention is embedded in another amino acid sequence.

As is demonstrated in the appended examples a two-hybrid screening assay has been developed in accordance with the present invention suitable for identifying cell cycle interacting proteins. Thus, in another aspect the present invention relates to a method for identifying and obtaining cycle interacting proteins comprising a two-hybrid screening assay wherein CDC2a as a bait and a cDNA library of vegetative plant tissue as prey are used. Preferably, said CDC2a is CDC2aAt. However, CDC2a from other organisms such as mammals may be employed as well.

The nucleic acid molecules encoding proteins or peptides identified to interact with the CDC2a in the above mentioned assay can be easily obtained and sequenced by methods known in the art; see also the appended examples. Therefore, the present

invention also relates to a DNA sequence encoding a cell cycle interacting protein obtainable by the method of the invention.

In a preferred embodiment the nucleic acid molecules according to the invention are RNA or DNA molecules, preferably cDNA, genomic DNA or synthetically synthesized DNA or RNA molecules. Preferably, the nucleic acid molecule of the invention is derived from a plant, preferably from *Arabidopsis thaliana*. As discussed above, the proteins encoded by the nucleic acid molecules identified according to the present invention in *Arabidopsis thaliana* show some homology to kinesins from several organisms. Corresponding proteins displaying similar properties should, therefore, be present in other plants as well. Nucleic acid molecules of the invention can be obtained, e.g., by hybridization of the above-described nucleic acid molecules with a (sample of) nucleic acid molecule(s) of any source. Nucleic acid molecules hybridizing with the above-described nucleic acid molecules can in general be derived from any organism, preferably plant possessing such molecules, preferably form monocotyledonous or dicotyledonous plants, in particular from any organism, preferably plants of interest in agriculture, horticulture or wood culture, such as crop plants, namely those of the family Poaceae, any starch producing plants, such as potato, maniok, leguminous plants, oil producing plants, such as oilseed rape, linenseed, etc., plants using polypeptide as storage substances, such as soybean, plants using sucrose as storage substance, such as sugar beet or sugar cane, trees, ornamental plants etc. Preferably, the nucleic acid molecules according to the invention are derived from *Arabidopsis thaliana*. Nucleic acid molecules hybridizing to the above-described nucleic acid molecules can be isolated, e.g., form libraries, such as cDNA or genomic libraries by techniques well known in the art. For example, hybridizing nucleic acid molecules can be identified and isolated by using the above-described nucleic acid molecules or fragments thereof or complements thereof as probes to screen libraries by hybridizing with said molecules according to standard techniques. Possible is also the isolation of such nucleic acid molecules by applying the polymerase chain reaction (PCR) using as primers oligonucleotides derived form the above-described nucleic acid molecules.

Nucleic acid molecules which hybridize with any of the aforementioned nucleic acid molecules also include fragments, derivatives and allelic variants of the above-described nucleic acid molecules that encode a cell cycle interacting protein or an immunologically or functional fragment thereof. Fragments are understood to be parts of nucleic acid molecules long enough to encode the described protein or a functional or immunologically active fragment thereof as defined above. Preferably, the functional fragment contains at least one of the phosphorylation sites and/or the Gln-rich domain at the N-terminus of the protein shown in Figure 1; see also Example 2.

The term "derivative" means in this context that the nucleotide sequence of these nucleic acid molecules differs from the sequences of the above-described nucleic acid molecules in one or more nucleotide positions and are highly homologous to said nucleic acid molecules. Homology is understood to refer to a sequence identity of at least 40 %, particularly an identity of at least 60 %, preferably more than 80 % and still more preferably more than 90 %. The term "substantially homologous" refers to a subject, for instance a nucleic acid, which is at least 50% identical in sequence to the reference when the entire ORF (open reading frame) is compared, where the sequence identity is preferably at least 70%, more preferably at least 80%, still more preferably at least 85%, especially more than about 90%, most preferably 95% or greater, particularly 98% or greater. The deviations from the sequences of the nucleic acid molecules described above can, for example, be the result of nucleotide substitution(s), deletion(s), addition(s), insertion(s) and/or recombination(s); see supra.

Homology further means that the respective nucleic acid molecules or encoded proteins are functionally and/or structurally equivalent. The nucleic acid molecules that are homologous to the nucleic acid molecules described above and that are derivatives of said nucleic acid molecules are, for example, variations of said nucleic acid molecules which represent modifications having the same biological function, in particular encoding proteins with the same or substantially the same biological function. They may be naturally occurring variations, such as sequences from other plant varieties or species, or mutations. These mutations may occur naturally or may

be obtained by mutagenesis techniques. The allelic variations may be naturally occurring allelic variants as well as synthetically produced or genetically engineered variants; see supra.

The proteins encoded by the various derivatives and variants of the above-described nucleic acid molecules share specific common characteristics, such as biological activity, molecular weight, immunological reactivity, conformation, etc., as well as physical properties, such as electrophoretic mobility, chromatographic behavior, sedimentation coefficients, pH optimum, temperature optimum, stability, solubility, spectroscopic properties, etc.

Examples of the different possible applications of the nucleic acid molecules according to the invention as well as molecules derived from them will be described in detail in the following.

Hence, in a further embodiment, the invention relates to nucleic acid molecules of at least 15 nucleotides in length hybridizing specifically with a nucleic acid molecule as described above or with a complementary strand thereof. Specific hybridization occurs preferably under stringent conditions and implies no or very little cross-hybridization with nucleotide sequences encoding no or substantially different proteins. Such nucleic acid molecules may be used as probes and/or for the control of gene expression. Nucleic acid probe technology is well known to those skilled in the art who will readily appreciate that such probes may vary in length. Preferred are nucleic acid probes of 16 to 35 nucleotides in length. Of course, it may also be appropriate to use nucleic acids of up to 100 and more nucleotides in length. The nucleic acid probes of the invention are useful for various applications. On the one hand, they may be used as PCR primers for amplification of nucleic acid sequences according to the invention. The design and use of said primers is known by the person skilled in the art. Preferably such amplification primers comprise a contiguous sequence of at least 6 nucleotides, in particular 13 nucleotides, preferably 15 to 25 nucleotides or more, identical or complementary to the nucleotide sequence depicted in SEQ ID NO: 1 or to a nucleotide sequence

encoding the amino acid sequence of SEQ ID NO: 2. Another application is the use as a hybridization probe to identify nucleic acid molecules hybridizing with a nucleic acid molecule of the invention by homology screening of genomic DNA or cDNA libraries. Nucleic acid molecules according to this preferred embodiment of the invention which are complementary to a nucleic acid molecule as described above may also be used for repression of expression of a cell cycle gene, for example due to an antisense or triple helix effect or for the construction of appropriate ribozymes (see, e.g., EP-A1 0 291 533, EP-A1 0 321 201, EP-A2 0 360 257) which specifically cleave the (pre)-mRNA of a gene comprising a nucleic acid molecule of the invention or part thereof. Selection of appropriate target sites and corresponding ribozymes can be done as described, for example, in Steinecke, Ribozymes, Methods in Cell Biology 50, Galbraith et al. eds Academic Press, Inc. (1995), 449-460. Furthermore, the person skilled in the art is well aware that it is also possible to label such a nucleic acid probe with an appropriate marker for specific applications, such as for the detection of the presence of a nucleic acid molecule of the invention in a sample derived from an organism, in particular plants.

The above described nucleic acid molecules may either be DNA or RNA or a hybrid thereof. Furthermore, said nucleic acid molecule may contain, for example, thioester bonds and/or nucleotide analogues, commonly used in oligonucleotide anti-sense approaches. Said modifications may be useful for the stabilization of the nucleic acid molecule against endo- and/or exonucleases in the cell. Said nucleic acid molecules may be transcribed by an appropriate vector containing a chimeric gene which allows for the transcription of said nucleic acid molecule in the cell.

Furthermore, the so-called "peptide nucleic acid" (PNA) technique can be used for the detection or inhibition of the expression of a nucleic acid molecule of the invention. For example, the binding of PNAs to complementary as well as various single stranded RNA and DNA nucleic acid molecules can be systematically investigated using thermal denaturation and BIAcore surface-interaction techniques (Jensen, Biochemistry 36 (1997), 5072-5077). Furthermore, the nucleic acid molecules described above as well as PNAs derived therefrom can be used for

detecting point mutations by hybridization with nucleic acids obtained from a sample with an affinity sensor, such as BIAcore; see Gotoh, Rinsho Byori 45 (1997), 224-228. Hybridization based DNA screening on peptide nucleic acids (PNA) oligomer arrays are described in the prior art, for example in Weiler, Nucleic Acids Research 25 (1997), 2792-2799. The synthesis of PNAs can be performed according to methods known in the art, for example, as described in Koch, J. Pept. Res. 49 (1997), 80-88; Finn, Nucleic Acids Research 24 (1996), 3357-3363. Further possible applications of such PNAs, for example as restriction enzymes or as templates for the synthesis of nucleic acid oligonucleotides are known to the person skilled in the art and are, for example, described in Veselkov, Nature 379 (1996), 214 and Bohler, Nature 376 (1995), 578-581.

The present invention also relates to vectors, particularly plasmids, cosmids, viruses, bacteriophages and other vectors used conventionally in genetic engineering that contain a nucleic acid molecule according to the invention. Methods which are well known to those skilled in the art can be used to construct various plasmids and vectors; see, for example, the techniques described in Sambrook, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory (1989) N.Y. and Ausubel, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1989). Alternatively, the nucleic acid molecules and vectors of the invention can be reconstituted into liposomes for delivery to target cells.

In a preferred embodiment the nucleic acid molecule present in the vector is linked to (a) control sequence(s) which allow the expression of the nucleic acid molecule in prokaryotic and/or eukaryotic cells.

The term "control sequence" refers to regulatory DNA sequences which are necessary to effect the expression of coding sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism. In prokaryotes, control sequences generally include promoter, ribosomal binding site, and terminators. In eukaryotes generally control sequences include promoters, terminators and, in some instances, enhancers, transactivators or transcription

factors. The term "control sequence" is intended to include, at a minimum, all components the presence of which are necessary for expression, and may also include additional advantageous components.

The term "operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences. In case the control sequence is a promoter, it is obvious for a skilled person that double-stranded nucleic acid is used.

Thus, the vector of the invention is preferably an expression vector. An "expression vector" is a construct that can be used to transform a selected host cell and provides for expression of a coding sequence in the selected host. Expression vectors can for instance be cloning vectors, binary vectors or integrating vectors. Expression comprises transcription of the nucleic acid molecule preferably into a translatable mRNA. Regulatory elements ensuring expression in prokaryotic and/or eukaryotic cells are well known to those skilled in the art. In the case of eukaryotic cells they comprise normally promoters ensuring initiation of transcription and optionally poly-A signals ensuring termination of transcription and stabilization of the transcript, for example, those of the 35S RNA from Cauliflower Mosaic Virus (CaMV). Other promoters commonly used are the polyubiquitin promoter, and the actin promoter for ubiquitous expression. The termination signals usually employed are from the Nopaline Synthase promoter or from the CAMV 35S promoter. A plant translational enhancer often used is the CAMV omega sequences, the inclusion of an intron (Intron-1 from the Shrunken gene of maize, for example) has been shown to increase expression levels by up to 100-fold. (Mait, Transgenic Research 6 (1997), 143-156; Ni, Plant Journal 7 (1995), 661-676). Additional regulatory elements may include transcriptional as well as translational enhancers. Possible regulatory elements permitting expression in prokaryotic host cells comprise, e.g., the P_L, *lac*, *trp* or *tac* promoter in *E. coli*, and examples of regulatory elements permitting expression in eukaryotic host cells are the *AOX1* or *GAL1* promoter in yeast or the CMV-, SV40- , RSV-promoter (Rous sarcoma virus), CMV-enhancer, SV40-enhancer or a globin intron in mammalian and other animal cells. In this context,

suitable expression vectors are known in the art such as Okayama-Berg cDNA expression vector pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3 (In-vitrogene), pSPORT1 (GIBCO BRL). Advantageously, the above-described vectors of the invention comprises a selectable and/or scorable marker. Selectable marker genes useful for the selection of transformed plant cells, callus, plant tissue and plants are well known to those skilled in the art and comprise, for example, antimetabolite resistance as the basis of selection for dhfr, which confers resistance to methotrexate (Reiss, Plant Physiol. (Life Sci. Adv.) 13 (1994), 143-149); npt, which confers resistance to the aminoglycosides neomycin, kanamycin and paromycin (Herrera-Estrella, EMBO J. 2 (1983), 987-995) and hygro, which confers resistance to hygromycin (Marsh, Gene 32 (1984), 481-485). Additional selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman, Proc. Natl. Acad. Sci. USA 85 (1988), 8047); mannose-6-phosphate isomerase which allows cells to utilize mannose (WO 94/20627) and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue, 1987, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed.) or deaminase from Aspergillus terreus which confers resistance to Blasticidin S (Tamura, Biosci. Biotechnol. Biochem. 59 (1995), 2336-2338).

Useful scorable marker are also known to those skilled in the art and are commercially available. Advantageously, said marker is a gene encoding luciferase (Giacomin, Pl. Sci. 116 (1996), 59-72; Scikantha, J. Bact. 178 (1996), 121), green fluorescent protein (Gerdes, FEBS Lett. 389 (1996), 44-47) or ß-glucuronidase (Jefferson, EMBO J. 6 (1987), 3901-3907). This embodiment is particularly useful for simple and rapid screening of cells, tissues and organisms containing a vector of the invention.

The present invention furthermore relates to host cells comprising a vector as described above or a nucleic acid molecule according to the invention wherein the nucleic acid molecule is foreign to the host cell.

By "foreign" it is meant that the nucleic acid molecule is either heterologous with respect to the host cell, this means derived from a cell or organism with a different genomic background, or is homologous with respect to the host cell but located in a different genomic environment than the naturally occurring counterpart of said nucleic acid molecule. This means that, if the nucleic acid molecule is homologous with respect to the host cell, it is not located in its natural location in the genome of said host cell, in particular it is surrounded by different genes. In this case the nucleic acid molecule may be either under the control of its own promoter or under the control of a heterologous promoter. The vector or nucleic acid molecule according to the invention which is present in the host cell may either be integrated into the genome of the host cell or it may be maintained in some form extrachromosomally. In this respect, it is also to be understood that the nucleic acid molecule of the invention can be used to restore or create a mutant gene via homologous recombination (Paszkowski (ed.), Homologous Recombination and Gene Silencing in Plants. Kluwer Academic Publishers (1994)).

The host cell can be any prokaryotic or eukaryotic cell, such as bacterial, insect, fungal, plant or animal cells. Preferred fungal cells are, for example, those of the genus Saccharomyces, in particular those of the species S. cerevisiae.

Another subject of the invention is a method for the preparation of cell cycle interacting proteins which comprises the cultivation of host cells according to the invention which, due to the presence of a vector or a nucleic acid molecule according to the invention, are able to express such a protein, under conditions which allow expression of the protein and recovering of the so-produced protein from the culture.

The term "expression" means the production of a protein or nucleotide sequence in the cell. However, said term also includes expression of the protein in a cell-free system. It includes transcription into an RNA product, post-transcriptional modification and/or translation to a protein product or polypeptide from a DNA encoding that product, as well as possible post-translational modifications. Depending on the specific constructs and conditions used, the protein may be

recovered from the cells, from the culture medium or from both. For the person skilled in the art it is well known that it is not only possible to express a native protein but also to express the protein as fusion polypeptides or to add signal sequences directing the protein to specific compartments of the host cell, e.g., ensuring secretion of the peptide into the culture medium, etc. Furthermore, such a protein and fragments thereof can be chemically synthesized and/or modified according to standard methods described, for example hereinbelow.

The terms "protein" and "polypeptide" used in this application are interchangeable. "Polypeptide" refers to a polymer of amino acids (amino acid sequence) and does not refer to a specific length of the molecule. Thus peptides and oligopeptides are included within the definition of polypeptide. This term does also refer to or include post-translational modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like. Included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

The present invention furthermore relates to proteins encoded by the nucleic acid molecules according to the invention or produced or obtained by the above-described methods, and to functional and/or immunologically active fragments of such cell cycle interacting proteins. The proteins and polypeptides of the present invention are not necessarily translated from a designated nucleic acid sequence; the polypeptides may be generated in any manner, including for example, chemical synthesis, or expression of a recombinant expression system, or isolation from a suitable viral system. The polypeptides may include one or more analogs of amino acids, phosphorylated amino acids or unnatural amino acids. Methods of inserting analogs of amino acids into a sequence are known in the art. The polypeptides may also include one or more labels, which are known to those skilled in the art. In this context, it is also understood that the proteins according to the invention may be further modified by conventional methods known in the art. By providing the proteins according to the present invention it is also possible to determine fragments which

retain biological activity, namely the mature, processed form. This allows the construction of chimeric proteins and peptides comprising an amino sequence derived from the protein of the invention, which is crucial for its binding activity and other functional amino acid sequences, e.g. GUS marker gene (Jefferson, EMBO J. 6 (1987), 3901-3907). The other functional amino acid sequences may be either physically linked by, e.g., chemical means to the proteins of the invention or may be fused by recombinant DNA techniques well known in the art.

The term "fragment of a sequence" or "part of a sequence" means a truncated sequence of the original sequence referred to. The truncated sequence (nucleic acid or protein sequence) can vary widely in length; the minimum size being a sequence of sufficient size to provide a sequence with at least a comparable function and/or activity of the original sequence referred to, while the maximum size is not critical. In some applications, the maximum size usually is not substantially greater than that required to provide the desired activity and/or function(s) of the original sequence. Typically, the truncated amino acid sequence will range from about 5 to about 60 amino acids in length. More typically, however, the sequence will be a maximum of about 50 amino acids in length, preferably a maximum of about 30 amino acids. It is usually desirable to select sequences of at least about 10, 12 or 15 amino acids, up to a maximum of about 20 or 25 amino acids.

Furthermore, folding simulations and computer redesign of structural motifs of the protein of the invention can be performed using appropriate computer programs (Olszewski, Proteins 25 (1996), 286-299; Hoffman, Comput. Appl. Biosci. 11 (1995), 675-679). Computer modeling of protein folding can be used for the conformational and energetic analysis of detailed peptide and protein models (Monge, J. Mol. Biol. 247 (1995), 995-1012; Renouf, Adv. Exp. Med. Biol. 376 (1995), 37-45). In particular, the appropriate programs can be used for the identification of interactive sites of the protein and cyclin dependent kinases, its receptor, its ligand or other interacting proteins by computer assistant searches for complementary peptide sequences (Fassina, Immunomethods 5 (1994), 114-120. Further appropriate computer systems for the design of protein and peptides are described in the prior art, for example in Berry, Biochem. Soc. Trans. 22 (1994), 1033-1036; Wodak, Ann. N. Y. Acad. Sci. 501 (1987), 1-13; Pabo, Biochemistry 25 (1986), 5987-5991. The results obtained from the

above-described computer analysis can be used for, e.g., the preparation of peptidomimetics of the protein of the invention or fragments thereof. Such pseudopeptide analogues of the natural amino acid sequence of the protein may very efficiently mimic the parent protein (Benkirane, J. Biol. Chem. 271 (1996), 33218-33224). For example, incorporation of easily available achiral Ω-amino acid residues into a protein of the invention or a fragment thereof results in the substitution of amide bonds by polymethylene units of an aliphatic chain, thereby providing a convenient strategy for constructing a peptidomimetic (Banerjee, Biopolymers 39 (1996), 769-777). Superactive peptidomimetic analogues of small peptide hormones in other systems are described in the prior art (Zhang, Biochem. Biophys. Res. Commun. 224 (1996), 327-331). Appropriate peptidomimetics of the protein of the present invention can also be identified by the synthesis of peptidomimetic combinatorial libraries through successive amide alkylation and testing the resulting compounds, e.g., for their binding and immunological properties. Methods for the generation and use of peptidomimetic combinatorial libraries are described in the prior art, for example in Ostresh, Methods in Enzymology 267 (1996), 220-234 and Dorner, Bioorg. Med. Chem. 4 (1996), 709-715.

Furthermore, a three-dimensional and/or crystallographic structure of the protein of the invention can be used for the design of peptidomimetic inhibitors of the biological activity of the protein of the invention (Rose, Biochemistry 35 (1996), 12933-12944; Rutenber, Bioorg. Med. Chem. 4 (1996), 1545-1558).

Furthermore, the present invention relates to antibodies specifically recognizing a cell cycle interacting protein according to the invention or parts, i.e. specific fragments or epitopes, of such a protein. The antibodies of the invention can be used to identify and isolate other cell cycle interacting proteins and genes in any organism, preferably plants. These antibodies can be monoclonal antibodies, polyclonal antibodies or synthetic antibodies as well as fragments of antibodies, such as Fab, Fv or scFv fragments etc. Monoclonal antibodies can be prepared, for example, by the techniques as originally described in Köhler and Milstein, Nature 256 (1975), 495, and Galfré, Meth. Enzymol. 73 (1981), 3, which comprise the fusion of mouse myeloma cells to spleen cells derived from immunized mammals. Furthermore, antibodies or

fragments thereof to the aforementioned peptides can be obtained by using methods which are described, e.g., in Harlow and Lane "Antibodies, A Laboratory Manual", CSH Press, Cold Spring Harbor, 1988. These antibodies can be used, for example, for the immunoprecipitation and immunolocalization of proteins according to the invention as well as for the monitoring of the synthesis of such proteins, for example, in recombinant organisms, and for the identification of compounds interacting with the protein according to the invention. For example, surface plasmon resonance as employed in the BIAcore system can be used to increase the efficiency of phage antibodies selections, yielding a high increment of affinity from a single library of phage antibodies which bind to an epitope of the protein of the invention (Schier, Human Antibodies Hybridomas 7 (1996), 97-105; Malmborg, J. Immunol. Methods 183 (1995), 7-13). In many cases, the binding phenomena of antibodies to antigens is equivalent to other ligand/anti-ligand binding.

Plant cell division can conceptually be influenced in three ways : (i) inhibiting or arresting cell division, (ii) maintaining, facilitating or stimulating cell division or (iii) uncoupling DNA synthesis from mitosis and cytokinesis. Modulation of the expression of a polypeptide encoded by a nucleotide sequence according to the invention has surprisingly an advantageous influence on plant cell division characteristics, in particular on the disruption of the expression levels of genes involved in G1/S and/or G2/M transition and as a result therof on the total make-up of the plant concerned or parts thereof. An example is that DNA synthesis or progression of DNA replication will be negatively influenced by elimination of specific substrates for a cyclin-dependent protein kinase complex.

The term "cyclin-dependent protein kinase complex" means the complex formed when a, preferably functional, cyclin associates with a, preferably, functional cyclin dependent kinase. Such complexes may be active in phosphorylating proteins and may or may not contain additional protein species.
The presence, absence or activity of a substrate for CDK in a plant cell is influenced by manipulation of the gene according to the invention. To analyse the industrial applicabilities of the invention, transformed plants can be made overproducing the

nucleotide sequence according to the invention. Such an overexpression of the new gene(s), proteins or inactivated variants thereof will either positively or negatively have an effect on cell division. Methods to modify the expression levels and/or the activity are known to persons skilled in the art and include for instance overexpression, co-suppression, the use of ribozymes, sense and anti-sense strategies, gene silencing approaches. "Sense strand" refers to the strand of a double-stranded DNA molecule that is homologous to a mRNA transcript thereof. The "anti-sense strand" contains an inverted sequence which is complementary to that of the "sense strand".

Hence, the nucleic acid molecules according to the invention are in particular useful for the genetic manipulation of plant cells in order to modify the characteristics of plants and to obtain plants with modified, preferably with improved or useful phenotypes. Similarly, the invention can also be used to modulate the cell division and the growth of cells, preferentially plant cells, in *in vitro* cultures.
Thus, the present invention provides for a method for the production of transgenic plants, plant cells or plant tissue comprising the introduction of a nucleic acid molecule or vector of the invention into the genome of said plant, plant cell or plant tissue.

For the expression of the nucleic acid molecules according to the invention in sense or antisense orientation in plant cells, the molecules are placed under the control of regulatory elements which ensure the expression in plant cells. These regulatory elements may be heterologous or homologous with respect to the nucleic acid molecule to be expressed as well with respect to the plant species to be transformed. In general, such regulatory elements comprise a promoter active in plant cells. To obtain expression in all tissues of a transgenic plant, preferably constitutive promoters are used, such as the 35 S promoter of CaMV (Odell, Nature 313 (1985), 810-812) or promoters of the polyubiquitin genes of maize (Christensen, Plant Mol. Biol. 18 (1982), 675-689). In order to achieve expression in specific tissues of a transgenic plant it is possible to use tissue specific promoters (see, e.g., Stockhaus, EMBO J. 8 (1989), 2245-2251). Known are also promoters which are specifically active in tubers of

potatoes or in seeds of different plants species, such as maize, Vicia, wheat, barley etc. Inducible promoters may be used in order to be able to exactly control expression. An example for inducible promoters are the promoters of genes encoding heat shock proteins. Also microspore-specific regulatory elements and their uses have been described (WO96/16182). Furthermore, the chemically inducible Test-system may be employed (Gatz, Mol. Gen. Genet. 227 (1991); 229-237). Further suitable promoters are known to the person skilled in the art and are described, e.g., in Ward (Plant Mol. Biol. 22 (1993), 361-366). The regulatory elements may further comprise transcriptional and/or translational enhancers functional in plants cells. Furthermore, the regulatory elements may include transcription termination signals, such as a poly-A signal, which lead to the addition of a poly A tail to the transcript which may improve its stability.

In the case that a nucleic acid molecule according to the invention is expressed in sense orientation it is in principle possible to modify the coding sequence in such a way that the protein is located in any desired compartment of the plant cell. These include the nucleus, endoplasmatic reticulum, the vacuole, the mitochondria, the plastids, the apoplast, the cytoplasm etc. Since CDC2, the interacting component of the protein of the invention excerts its effects in the cytoplasm and/or nucleus, corresponding signal sequences are preferred to direct the protein of the invention in the same compartment. Methods how to carry out this modifications and signal sequences ensuring localization in a desired compartment are well known to the person skilled in the art.

Methods for the introduction of foreign DNA into plants are also well known in the art. These include, for example, the transformation of plant cells or tissues with T-DNA using Agrobacterium tumefaciens or Agrobacterium rhizogenes, the fusion of protoplasts, direct gene transfer (see, e.g., EP-A 164 575), injection, electroporation, biolistic methods like particle bombardment, pollen-mediated transformation, plant RNA virus-mediated transformation, liposome-mediated transformation, transformation using wounded or enzyme-degraded immature embryos, or wounded or enzyme-degraded embryogenic callus and other methods known in the art. The

vectors used in the method of the invention may contain further functional elements, for example "left border"- and "right border"-sequences of the T-DNA of Agrobacterium which allow for stably integration into the plant genome. Furthermore, methods and vectors are known to the person skilled in the art which permit the generation of marker free transgenic plants, i.e. the selectable or scorable marker gene is lost at a certain stage of plant development or plant breeding. This can be achieved by, for example cotransformation (Lyznik, Plant Mol. Biol. 13 (1989), 151-161; Peng, Plant Mol. Biol. 27 (1995), 91-104) and/or by using systems which utilize enzymes capable of promoting homologous recombination in plants (see, e.g., WO97/08331; Bayley, Plant Mol. Biol. 18 (1992), 353-361); Lloyd, Mol. Gen. Genet. 242 (1994), 653-657; Maeser, Mol. Gen. Genet. 230 (1991), 170-176; Onouchi, Nucl. Acids Res. 19 (1991), 6373-6378). Methods for the preparation of appropriate vectors are described by, e.g., Sambrook (Molecular Cloning; A Laboratory Manual, 2nd Edition (1989), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Suitable strains of Agrobacterium tumefaciens and vectors as well as transformation of Agrobacteria and appropriate growth and selection media are well known to those skilled in the art and are described in the prior art (GV3101 (pMK90RK), Koncz, Mol. Gen. Genet. 204 (1986), 383-396; C58C1 (pGV 3850kan), Deblaere, Nucl. Acid Res. 13 (1985), 4777; Bevan, Nucleic. Acid Res. 12(1984), 8711; Koncz, Proc. Natl. Acad. Sci. USA 86 (1989), 8467-8471; Koncz, Plant Mol. Biol. 20 (1992), 963-976; Koncz, Specialized vectors for gene tagging and expression studies. In: Plant Molecular Biology Manual Vol 2, Gelvin and Schilperoort (Eds.), Dordrecht, The Netherlands: Kluwer Academic Publ. (1994), 1-22; EP-A-120 516; Hoekema: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasserdam (1985), Chapter V, Fraley, Crit. Rev. Plant. Sci., 4, 1-46; An, EMBO J. 4 (1985), 277-287). Although the use of Agrobacterium tumefaciens is preferred in the method of the invention, other Agrobacterium strains, such as Agrobacterium rhizogenes, may be used, for example if a phenotype conferred by said strain is desired.

Methods for the transformation using biolistic methods are well known to the person skilled in the art; see, e.g., Wan, Plant Physiol. 104 (1994), 37-48; Vasil, Bio/Technology 11 (1993), 1553-1558 and Christou (1996) Trends in Plant Science

1, 423-431. Microinjection can be performed as described in Potrykus and Spangenberg (eds.), Gene Transfer To Plants. Springer Verlag, Berlin, NY (1995).

The transformation of most dicotyledonous plants is possible with the methods described above. But also for the transformation of monocotyledonous plants several successful transformation techniques have been developed. These include the transformation using biolistic methods as, e.g., described above as well as protoplast transformation, electroporation of partially permeabilized cells, introduction of DNA using glass fibers, etc.

The term "transformation" as used herein, refers to the transfer of an exogenous polynucleotide into a host cell, irrespective of the method used for the transfer. The polynucleotide may be transiently or stably introduced into the host cell and may be maintained non-integrated, for example, as a plasmid, or alternatively, may be integrated into the host genome. The resulting transformed plant cell can then be used to regenerate a transformed plant in a manner known by a skilled person.

In general, the plants which can be modified according to the invention and which either show overexpression of a protein according to the invention or a reduction of the synthesis of such a protein can be derived from any desired plant species. They can be monocotyledonous plants or dicotyledonous plants, preferably they belong to plant species of interest in agriculture, wood culture or horticulture interest, such as crop plants (e.g. maize, rice, barley, wheat, rye, oats etc.), potatoes, oil producing plants (e.g. oilseed rape, sunflower, pea nut, soy bean, etc.), cotton, sugar beet, sugar cane, leguminous plants (e.g. beans, peas etc.), wood producing plants, preferably trees, etc.

Thus, the present invention relates also to transgenic plant cells which contain stably integrated into the genome a nucleic acid molecule according to the invention linked to regulatory elements which allow for expression of the nucleic acid molecule in plant cells and wherein the nucleic acid molecule is foreign to the transgenic plant cell. For the meaning of foreign; see supra.

The presence and expression of the nucleic acid molecule in the transgenic plant cells leads to the synthesis of a cell cycle interacting protein and leads to physiological and phenotypic changes in plants containing such cells.

Thus, the present invention also relates to transgenic plants and plant tissue comprising transgenic plant cells according to the invention. Due to the (over)expression of a cell cycle interacting protein of the invention, e.g., at developmental stages and/or in plant tissue in which they do not naturally occur these transgenic plants may show various physiological, developmental and/or morphological modifications in comparison to wild-type plants. For example, these transgenic plants may display an altered cell elongation.

Therefore, part of this invention is the use of plant cell cycle genes and/or plant cell cycle proteins to modulate plant cell division and/or growth in plant cells, plant tissues, plant organs and/or whole plants. To the scope of the invention also belongs a method to influence the activity of cyclin-dependent protein kinase in a plant cell by transforming the plant cell with a nucleic acid molecule according to the invention and/or manipulation of the expression of said molecule. More in particular using a nucleic acid molecule according to the invention, the disruption of plant cell division can be accomplished by interfering in the expression of a substrate for cyclin-dependent protein kinase. The latter goal may be achieved, for example, with methods for reducing the amount of active cell cycle interacting proteins.

Hence, the invention also relates to a transgenic plant cell which contains (stably integrated into the genome) a nucleic acid molecule according to the invention or part thereof, wherein the transcription and/or expression of the nucleic acid molecule or part thereof leads to reduction of the synthesis of a cell cycle interacting protein.

In a preferred embodiment, the reduction is achieved by an anti-sense, sense, ribozyme, co-suppression and/or dominant mutant effect.

"Antisense" and "antisense nucleotides" means DNA or RNA constructs which block the expression of the naturally occurring gene product. For example, in the present

invention use of a DNA construct that produces "th65 antisense RNA" blocks the expression of "th65" by destroying or inactivating "th65 mRNA".

The provision of the nucleic acid molecules according to the invention opens up the possibility to produce transgenic plant cells with a reduced level of the protein as described above and, thus, with a defect in the accumulation of a cell cycle interacting protein. Techniques how to achieve this are well known to the person skilled in the art. These include, for example, the expression of antisense-RNA, ribozymes, of molecules which combine antisense and ribozyme functions and/or of molecules which provide for a co-suppression effect; see also supra. When using the antisense approach for reduction of the amount of cell cycle interacting proteins in plant cells, the nucleic acid molecule encoding the antisense-RNA is preferably of homologous origin with respect to the plant species used for transformation. However, it is also possible to use nucleic acid molecules which display a high degree of homology to endogenously occurring nucleic acid molecules encoding a cell cycle interacting protein. In this case the homology is preferably higher than 80%, particularly higher than 90% and still more preferably higher than 95%.

The reduction of the synthesis of a protein according to the invention in the transgenic plant cells can result in an alteration in, e.g., cell division. In transgenic plants comprising such cells this can lead to various physiological, developmental and/or morphological changes.

Thus, the present invention also relates to transgenic plants comprising the above-described transgenic plant cells. These may show, for example, reduced growth characteristics.

The present invention also relates to cultured plant tissues comprising transgenic plant cells as described above which either show overexpression of a protein according to the invention or a reduction in synthesis of such a protein.

Any transformed plant obtained according to the invention can be used in a conventional breeding scheme or in *in vitro* plant propagation to produce more transformed plants with the same characteristics and/or can be used to introduce the

same characteristic in other varieties of the same or related species. Such plants are also part of the invention. Seeds obtained from the transformed plants genetically also contain the same characteristic and are part of the invention. As mentioned before, the present invention is in principle applicable to any plant and crop that can be transformed with any of the transformation method known to those skilled in the art and includes for instance corn, wheat, barley, rice, oilseed crops, cotton, tree species, sugar beet, cassava, tomato, potato, numerous other vegetables, fruits.

In yet another aspect, the invention also relates to harvestable parts and to propagation material of the transgenic plants according to the invention which either contain transgenic plant cells expressing a nucleic acid molecule according to the invention or which contain cells which show a reduced level of the described protein. Harvestable parts can be in principle any useful parts of a plant, for example, flowers, pollen, seedlings, tubers, leaves, stems, fruit, seeds, roots etc. Propagation material includes, for example, seeds, fruits, cuttings, seedlings, tubers, rootstocks etc.

The present invention further relates to a method for identifying and obtaining an activator or inhibitor of cell cycle proteins comprising the steps of:

(a) combining a compound to be screened with a reaction mixture containing the protein of the invention and a readout system capable of interacting with the protein under suitable conditions;

(b) maintaining said reaction mixture in the presence of the compound or a sample comprising a plurality of compounds under conditions which permit interaction of the protein with said readout system;

(c) identifying or verifying a sample and compound, respectively, which leads to suppression or activation of the readout system.

The term "read out system" in context with the present invention means a DNA sequence which upon transcription and/or expression in a cell, tissue or organism provides for a scorable and/or selectable phenotype. Such read out systems are well known to those skilled in the art and comprise, for example, recombinant DNA molecules and marker genes as described above and in the appended example.

The term "plurality of compounds" in a method of the invention is to be understood as a plurality of substances which may or may not be identical.

Said compound or plurality of compounds may be comprised in, for example, samples, e.g., cell extracts from, e.g., plants, animals or microorganisms. Furthermore, said compound(s) may be known in the art but hitherto not known to be capable of suppressing or activating cell cycle interacting proteins. The reaction mixture may be a cell free extract or may comprise a cell or tissue culture. Suitable set ups for the method of the invention are known to the person skilled in the art and are, for example, generally described in Alberts et al., Molecular Biology of the Cell, third edition (1994), in particular Chapter 17. The plurality of compounds may be, e.g., added to the reaction mixture, culture medium or injected into the cell.

If a sample containing a compound or a plurality of compounds is identified in the method of the invention, then it is either possible to isolate the compound from the original sample identified as containing the compound capable of suppressing or activating cell cycle interacting proteins, or one can further subdivide the original sample, for example, if it consists of a plurality of different compounds, so as to reduce the number of different substances per sample and repeat the method with the subdivisions of the original sample. Depending on the complexity of the samples, the steps described above can be performed several times, preferably until the sample identified according to the method of the invention only comprises a limited number of or only one substance(s). Preferably said sample comprises substances of similar chemical and/or physical properties, and most preferably said substances are identical. Preferably, the compound identified according to the above described method or its derivative is further formulated in a form suitable for the application in plant breeding or plant cell and tissue culture.

The compounds which can be tested and identified according to a method of the invention may be expression libraries, e.g., cDNA expression libraries, peptides, proteins, nucleic acids, antibodies, small organic compounds, hormones, peptidomimetics, PNAs or the like (Milner, Nature Medicine 1 (1995), 879-880; Hupp, Cell 83 (1995), 237-245; Gibbs, Cell 79 (1994), 193-198 and references cited supra). Furthermore, genes encoding a putative regulator of cell cycle interacting

gene and/or which excert their effects up- or downstream the cell cycle interacting protein of the invention may be identified using, for example, insertion mutagenesis using, for example, gene targeting vectors known in the art (see, e.g., Hayashi, Science 258 (1992), 1350-1353; Fritze and Walden, Gene activation by T-DNA tagging. In *Methods in Molecular* biology 44 (Gartland, K.M.A. and Davey, M.R., eds). Totowa: Human Press (1995), 281-294) or transposon tagging (Chandlee, Physiologia Plantarum 78 (1990), 105-115). Said compounds can also be functional derivatives or analogues of known inhibitors or activators. Methods for the preparation of chemical derivatives and analogues are well known to those skilled in the art and are described in, for example, Beilstein, Handbook of Organic Chemistry, Springer edition New York Inc., 175 Fifth Avenue, New York, N.Y. 10010 U.S.A. and Organic Synthesis, Wiley, New York, USA. Furthermore, said derivatives and analogues can be tested for their effects according to methods known in the art. Furthermore, peptidomimetics and/or computer aided design of appropriate derivatives and analogues can be used, for example, according to the methods described above. The cell or tissue that may be employed in the method of the invention preferably is a host cell, plant cell or plant tissue of the invention described in the embodiments hereinbefore.

Determining whether a compound is capable of suppressing or activating cell cycle interacting proteins can be done, for example, by monitoring DNA duplication and cell division. It can further be done by monitoring the phenotypic characteristics of the cell of the invention contacted with the compounds and compare it to that of wild-type plants. In an additional embodiment, said characteristics may be compared to that of a cell contacted with a compound which is either known to be capable or incapable of suppressing or activating cell cycle interacting proteins.

The inhibitor or activator identified by the above-described method may prove useful as a herbicide, pesticide and/or as a plant growth regulator. Thus, in a further embodiment the invention relates to a compound obtained or identified according to the method of the invention said compound being an activator of cell cycle interacting proteins or an inhibitor of cell cycle interacting proteins. The above-

described compounds include, for example, cell cycle kinase inhibitors. "Cell-cycle kinase inhibitor" (CKI) is a protein which inhibit CDK/cyclin activity and is produced and/or activated when further cell division has to be temporarily or continuously prevented.

Such useful compounds can be for example transacting factors which bind to the cell cycle interacting protein of the invention. Identification of transacting factors can be carried out using standard methods in the art (see, e.g., Sambrook, supra, and Ausubel, supra). To determine whether a protein binds to the protein of the invention, standard native gel-shift analyses can be carried out. In order to identify a transacting factor which binds to the protein of the invention, the protein of the invention can be used as an affinity reagent in standard protein purification methods, or as a probe for screening an expression library. Once the transacting factor is identified, modulation of its binding to the cell cycle interacting protein of the invention can be pursued, beginning with, for example, screening for inhibitors against the binding of the transacting factor to the protein of the present invention. Activation or repression of cell cycle interacting proteins could then be achieved in plants by applying of the transacting factor (or its inhibitor) or the gene encoding it, e.g. in a vector for transgenic plants. In addition, if the active form of the transacting factor is a dimer, dominant-negative mutants of the transacting factor could be made in order to inhibit its activity. Furthermore, upon identification of the transacting factor, further components in the pathway leading to activation (e.g. signal transduction) or repression of a gene involved in the control of cell cycle then can be identified. Modulation of the activities of these components can then be pursued, in order to develop additional drugs and methods for modulating the cell cycle in animals and plants.

The invention also relates to a diagnostic composition comprising at least one of the aforementioned nucleic acid molecules, vectors, proteins, antibodies or compounds and optionally suitable means for detection.
Said diagnostic compositions may be used for methods for detecting expression of cell cycle interacting proteins by detecting the presence of the corresponding mRNA

which comprises isolation of mRNA from a cell and contacting the mRNA so obtained with a probe comprising a nucleic acid probe as described above under hybridizing conditions, detecting the presence of mRNA hybridized to the probe, and thereby detecting the expression of the protein in the cell. Further methods of detecting the presence of a protein according to the present invention comprises immunotechniques well known in the art, for example enzyme linked immunosorbent assay. Furthermore, it is possible to use the nucleic acid molecules according to the invention as molecular markers in plant breeding.

The person skilled in the art can use proteins according to the invention from other organisms such as yeast and animals to influence cell division progression in those other organisms such as mammals or insects. In a preferred embodiment one or more DNA sequences, vectors or proteins of the invention or the above-described antibody or compound are, for instance, used to specifically interfere in the disruption of the expression levels of genes involved in G1/S and/or G2/M transition in the cell cycle process in transformed plants, particularly :

- in the complete plant
- in selected plant organs, tissues or cell types
- under specific environmental conditions, including abiotic stress such as cold, heat, drought or salt stress or biotic stress such as pathogen attack
- during specific developmental stages.

Another aspect of the current invention is that one or more DNA sequences, vectors or proteins of the invention or the above-described antibody or compound can be used to modulate, for instance, endoreduplication in storage cells, storage tissues and/or storage organs of plants or parts thereof. The term "endoreduplication" means recurrent DNA replication without consequent mitosis and cytokinesis. Preferred target storage organs and parts thereof for the modulation of endoreduplication are, for instance, seeds (such as from cereals, oilseed crops), roots (such as in sugar beet), tubers (such as in potato) and fruits (such as in vegetables and fruit species). Furthermore it is expected that increased endoreduplication in storage organs and parts thereof correlates with enhanced storage capacity and as such with improved yield. In yet another embodiment of the

invention, a plant with modulated endoreduplication in the whole plant or parts thereof can be obtained from a single plant cell by transforming the cell, in a manner known to the skilled person, with the above-described means.

In view of the foregoing, the present invention also relates to the use of a DNA sequence, vector, protein, antibody or compound of the invention for modulating plant cell cycle, plant cell division and/or growth, for influencing the activity of cyclin-dependent protein kinase in a plant cell, for disrupting plant cell division by influencing the presence or absence or by interfering in the expression of a substrate for cyclin-dependent protein kinase, for influencing cell division progression in a host as defined above or for use in a screening method for the identification of inhibitors or activators of cell cycle proteins. Beside the above described possibilities to use the nucleic acid molecules according to the invention for the genetic engineering of plants with modified characteristics and their use to identify homologous molecules, the described nucleic acid molecules may also be used for several other applications, for example, for the identification of nucleic acid molecules which encode proteins which interact with the cell cycle proteins described above. This can be achieved by assays well known in the art such as those described above and also included, for example, as described in Scofield (Science 274 (1996), 2063-2065) by use of the so-called yeast "two-hybrid system"; see also the appended examples. In this system the protein encoded by the nucleic acid molecules according to the invention or a smaller part thereof is linked to the DNA-binding domain of the GAL4 transcription factor. A yeast strain expressing this fusion protein and comprising a lacZ reporter gene driven by an appropriate promoter, which is recognized by the GAL4 transcription factor, is transformed with a library of cDNAs which will express plant proteins or peptides thereof fused to an activation domain. Thus, if a peptide encoded by one of the cDNAs is able to interact with the fusion peptide comprising a peptide of a protein of the invention, the complex is able to direct expression of the reporter gene. In this way the nucleic acid molecules according to the invention and the encoded peptide can be used to identify peptides and proteins interacting with cell cycle interacting proteins. It is apparent to the person skilled in the art that this and similar systems may then further be exploited for the identification of inhibitors of the binding of the interacting proteins.

Other methods for identifying compounds which interact with the proteins according to the invention or nucleic acid molecules encoding such molecules are, for example, the in vitro screening with the phage display system as well as filter binding assays or "real time" measuring of interaction using, for example, the BIAcore apparatus (Pharmacia); see references cited supra.

These and other embodiments are disclosed and encompassed by the description and examples of the present invention. Further literature concerning any one of the methods, uses and compounds to be employed in accordance with the present invention may be retrieved from public libraries, using for example electronic devices. For example the public database "Medline" may be utilized which is available on the Internet, for example under http://www.ncbi.nlm.nih.gov/PubMed/medline.html. Further databases and addresses, such as http://www.ncbi.nlm.nih.gov/, http://www.infobiogen.fr/, http://www.fmi.ch/biology/research_tools.html, http://www.tigr.org/, are known to the person skilled in the art and can also be obtained using, e.g., http://www.lycos.com. An overview of patent information in biotechnology and a survey of relevant sources of patent information useful for retrospective searching and for current awareness is given in Berks, TIBTECH 12 (1994), 352-364.

The present invention is further described by reference to the following non-limiting figures and examples.

The Figures show:

Figure 1. Amino acid sequence encoded by the *th65* clone. TPNK, SPGR, and SPVR (in bold type and underlined) in the C-terminal part of the sequence are CDK consensus phosphorylation sites. In the N-terminal region of the sequence the amino acid residue Q (underlined) is repeatedly present.

The Examples illustrate the invention:

Unless stated otherwise in the examples, all recombinant DNA techniques are performed according to protocols as described in Sambrook et al. (1989), Molecular Cloning : A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY or in Volumes 1 and 2 of Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Standard materials and methods for plant molecular work are described in Plant Molecular Biology Labfase (1993) by R.D.D. Croy, jointly published by BIOS Scientific Publications Ltd (UK) and Blackwell Scientific Publications (UK).

Example 1: Identification of a cell cycle interacting protein

The used vectors and strains were supplied with the Matchmaker Two-Hybrid System (Clontech, Palo Alto, CA). Baits using CDC2aAt were constructed by inserting PCR fragments into the pGBT9 vector. The PCR fragments were created from the cDNAs by using primers to incorporate *Eco*RI restriction enzyme sites. For *CDC2aAt*, the primers 5'-CGAGATCTGAATTCATGGATCAGTA-3' (SEQ ID NO: 3) and 5'-CGAGATCTGAATTCCTAAGGCATGCC-3' (SEQ ID NO: 4) were used. For *CDC2bAt* the primers 5'-CGGATCCGAATTCATGGAGAACGAG-3' (SEQ ID NO: 5) and 5'-CGGATCCGAATTCTCAGAACTGAGA-3' (SEQ ID NO: 6) were used. The PCR fragments were cut with *Eco*RI and cloned into the *Eco*RI site of pGBT9, resulting in the plasmids pGBTCDC2A and pGBTCDC2B. The GAL4 activation domain cDNA fusion library of 3-week-old vegetative *Arabidopsis* plants was obtained from Clontech. For the screening, a 1-liter culture of the *Saccharomyces cerevisiae* strain HF7c (*MAT*$_a$ *ura3-52 his3-200 ade2-101 lys2-801 trp1-901 leu2-3,112 gal4-542 gal80-538 LYS2::GAL1*$_{UAS}$*-GAL1*$_{TATA}$*-HIS3 URA3::GAL4*$_{17mers(3x)}$*-CyC1*$_{TATA}$*-LacZ*) was cotransformed with 590 µg pGBTCDC2A, 1100 µg DNA of the library, and 40 mg salmon sperm carrier DNA using the lithium acetate method (Gietz *et al.*, 1992, Nucleic Acids Research, 20, pg.1425). To estimate the number of independent cotransformants, 1/1000 of the transformation mix was plated on Leu$^-$ and Trp$^-$

medium. The rest of the transformation mix was plated on medium to select for histidine prototrophy (Trp^-, Leu^-, His^-). After 6 days of growth at 30°C, the colonies larger than 2 mm were streaked on histidine-lacking medium supplemented with 10 mM 3-amino-1,2,4-triazole (Sigma, St. Louis, MO). Colonies capable of growing under these conditions were tested for β-galactosidase activity. The activation domain plasmids were isolated from the His^+ and $LacZ^+$ colonies. The pGAD10 inserts were PCR amplified using the primers 5'-ATACCACTACAATGGATG-3' (SEQ ID NO: 7) and 5'-AGTTGAAGTGAACTTGCGGG-3' (SEQ ID NO: 8). PCR fragments were digested with *Alu*I and fractionized on a 2% agarose gel. Plasmids whose PCR product gave rise to distinct restriction patterns were electroporated into *Escherichia coli* XL1-Blue, and the DNA sequence of the inserts was determined. Extracted DNA was also used to retransform HF7c to test the specificity of the interaction.

Example 2: Characterisation of the novel cell cycle gene

Above-mentioned two-hybrid screening was performed using as bait a fusion protein between the GAL4 DNA-binding domain and CDC2aAt. For the screening a GAL4 activation domain cDNA fusion library was used, constructed from 3-week-old vegetative tissue of *Arabidopsis thaliana*. After sequential selection rounds, an interesting clone encoding a CDC2aAt-specific interacting protein was identified being designated as *th65*, and appears to contain a 394-amino-acid-long open reading frame (figure 1) corresponding to SEQ ID NO 2, which had no significant homology with any protein in data bases. The N-terminus contained a region rich in glutamine residues (SEQ ID NO 2.). Gln-rich domains are often part of the transcriptional activation domain of DNA binding factors (Mitchell and Tjian, 1989, Science, 245, 371-378) and have also been shown to be involved in protein-protein interactions (Bao *et al.*, 1996, PNAS, 93, 5037-5042). The th65 open reading frame also contains three consensus CDK phosphorylation sites. The identification of th65 as a CDC2aAt-associated protein and the presence of these phosphorylation sites indicates that the th65 protein is a substrate for CDKs.

Example 3: Generation f transgenic plants with alt red cell cycle

A genomic clone of the *TH65* gene was obtained by standard procedures and entirely sequenced. The full length coding region was subsequently cloned in sense and antisense orientation in the binary vector PGSV4 (Herouart et al., 1994, Plant Physiol. 104, p 873-886) under the control of the constitutive *CaMV 35S* promoter. Additionally, a construct containing substitutions of the consensus CDK phosphorylation sites into non-phosphorylatable sites was constructed by point-mutagenesis and cloned in PGSV4 under the control of the *CaMV 35S* promoter. The obtained binary vectors were transformed to *Agrobacterium tumefaciens*. These strains were used to transform *Nicotiana tabacum* cv. Petit havana using the leaf disk protocol (Horsh et al., 1985, Science 227, p 1229-1231) and *Arabidopsis thaliana* using the root transformation protocol (Valvekens et al., 1988, PNAS 85, p 5536-5540).

Claims

1. A DNA sequence encoding a cell cycle interacting protein or encoding an immunologically active and/or functional fragment of such a protein, selected from the group consisting of:

   (a) DNA sequences comprising a nucleotide sequence encoding a protein comprising the amino acid sequence as given in SEQ ID NO: 2;

   (b) DNA sequences comprising a nucleotide sequence as given in SEQ ID NO: 1;

   (c) DNA sequences hybridizing with the complementary strand of a DNA sequence as defined in (a) or (b) and encoding an amino acid sequence which is at least 80% identical to the amino acid sequence encoded by the DNA sequence of (a) or (b);

   (d) DNA sequences, the nucleotide sequence of which is degenerated as a result of the genetic code to a nucleotide sequence of a DNA sequence as defined in any one of (a) to (c); and

   (e) DNA sequences encoding a fragment of a protein encoded by a DNA sequence of any one of (a) to (d).

2. A method for identifying and obtaining cycle interacting proteins comprising a two-hybrid screening assay wherein CDC2a as a bait and a cDNA library of vegetative plant tissue as prey are used.

3. The method of claim 2, wherein said CDC2a is CDC2aAt.

4. A DNA sequence encoding a cell cycle interacting protein obtainable by the method of claim 2 or 3.

5. A nucleic acid molecule of at least 15 nucleotides in length hybridizing specifically with a DNA sequence of claim 1 or 4 or with a complementary strand thereof.

6. A vector comprising a DNA sequence of claim 1 or 4.

7. The vector of claim 6 which is an expression vector wherein the DNA sequence is operatively linked to one or more control sequences allowing the expression in prokaryotic and/or eukaryotic host cells.

8. A host cell containing a vector of claim 6 or 7 or a DNA sequence of claim 1 or 4.

9. The host cell of claim 8 which is a bacterial, insect, fungal, plant or animal cell.

10. A method for the production of a cell cycle interacting protein or an immunologically active or functional fragment thereof comprising culturing a host cell of claim 8 or 9 under conditions allowing the expression of the protein and recovering the produced protein from the culture.

11. A cell cycle interacting protein or an immunologically active or functional fragment thereof encodable by a DNA sequence of claim 1 or 4 or obtainable by the method of claim 2, 3 or 10.

12. An antibody specifically recognizing the protein of claim 11 or a fragment or epitope thereof.

13. A method for the production of transgenic plants, plant cells or plant tissue comprising the introduction of a DNA sequence of claim 1, 4 or 5 or a vector of claim 6 or 7 into the genome of said plant, plant cell or plant tissue.

14. The method of claim 13 further comprising regenerating a plant from said plant tissue or plant cell.

15. A transgenic plant cell comprising a DNA sequence of claim 1 or 4 which is operably linked to regulatory elements allowing transcription and/or expression of the DNA sequence in plant cells or obtainable according to the method of claim 13 or 14.

16. The transgenic plant cell of claim 15 wherein said DNA sequence or said vector is stably integrated into the genome of the plant cell.

17. A transgenic plant or a plant tissue comprising plant cells of claim 15 or 16.

18. The transgenic plant of claim 17 in which plant cell division and/or growth is altered.

19. A transgenic plant cell which contains stably integrated into the genome a DNA sequence of claim 1, 4 or 5 or part thereof or obtainable according to the method of claim 13 or 14, wherein the transcription and/or expression of the DNA sequence or part thereof leads to reduction of the synthesis of the protein of claim 11 in the cells.

20. The plant cell of claim 19, wherein the reduction is achieved by an antisense, sense, ribozyme, co-suppression and/or dominant mutant effect.

21. A transgenic plant or plant tissue comprising the plant cells of claim 19 or 20.

22. The transgenic plant of claim 21 which displays a deficiency in plant cell division and/or growth.

23. Harvestable parts or propagation material of plants of any one of claims 17, 18, 21 or 22 comprising plant cells of claim 15, 16, 19 or 20.

24. A method for identifying and obtaining an activator or inhibitor of cell cycle proteins comprising the steps of:

(a) combining a compound to be screened with a reaction mixture containing the protein of claim 11 and a readout system capable of interacting with the protein under suitable conditions;

(b) maintaining said reaction mixture in the presence of the compound or a sample comprising a plurality of compounds under conditions which permit interaction of the protein with said readout system;

(d) identifying or verifying a sample and compound, respectively, which leads to suppression or activation of the readout system.

25. A method of producing a plant herbicide comprising the steps of the method of claim 24 and formulating the compound obtained or identified in step (c) or a derivative thereof in a form suitable for the application in agriculture or plant cell and tissue culture.

26. A compound obtained or identified by the method of claim 24, which is an activator or inhibitor of cell cycle interacting proteins.

27. A diagnostic composition comprising a DNA sequence of claim 1, 4 or 5, a vector of claim 6 or 7, a protein of claim 11, an antibody of claim 12, or the compound of claim 26, and optionally suitable means for detection.

28. Use of a DNA sequence of claim 1, 4 or 5, the vector of claim 6 or 7, the protein of claim 11, the antibody of claim 12 or the compound of claim 26 for modulating plant cell cycle, plant cell division and/or growth, for influencing the activity of cyclin-dependent protein kinase in a plant cell, for disrupting plant cell division by influencing the presence or absence or by interfering in the expression of a substrate for cyclin-dependent protein kinase for influencing cell division progression in a host as defined in claim 9 or for use in a screening method for the identification of inhibitors or activators of cell cycle proteins.

29. Use of the compound of claim 26 as growth regulator and/or herbicide.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: CropDesign NV
(B) STREET: TechnologiePark Zwijnaarde 3
(C) CITY: Gent
(D) STATE: none
(E) COUNTRY: Belgium
(F) POSTAL CODE (ZIP): 9052

(ii) TITLE OF INVENTION: Method and means for modulating plant cell cycle proteins and their use in controlling plant cell growth

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1184 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..1184

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
CG GCA AGT GAT GCT CGG AAG GAG CTG TTG GAG AAG GAG AGA GAA AAT          47
   Ala Ser Asp Ala Arg Lys Glu Leu Leu Glu Lys Glu Arg Glu Asn
     1               5                  10                  15

CAG AAT CTG AAA CAA GAG GTT GTG GGC TTA AAA AAA GCT CTT AAA GAT          95
Gln Asn Leu Lys Gln Glu Val Val Gly Leu Lys Lys Ala Leu Lys Asp
                20                  25                  30
```

```
GCA AAT GAC CAG TGT GTA TTA CTC TAC AGT GAA GTG CAG AGA GCG TGG        143
Ala Asn Asp Gln Cys Val Leu Leu Tyr Ser Glu Val Gln Arg Ala Trp
                35                  40                  45

AAA GTT TCA TTT ACA TTG CAA TCA GAT TTA AAG TCA GAG AAT ATT ATG        191
Lys Val Ser Phe Thr Leu Gln Ser Asp Leu Lys Ser Glu Asn Ile Met
        50                  55                  60

CTT GTA GAC AAA CAT AGA CTA GAG AAG GAG CAG AAT TCT CAG TTA AGG        239
Leu Val Asp Lys His Arg Leu Glu Lys Glu Gln Asn Ser Gln Leu Arg
    65                  70                  75

AAT CAA ATA GCT CAA TTT TTA CAG TTG GAT CAG GAA CAG AAG CTG CAA        287
Asn Gln Ile Ala Gln Phe Leu Gln Leu Asp Gln Glu Gln Lys Leu Gln
80                  85                  90                  95

ATG CAA CAA CAA GAT TCC GCC ATT CAA AAT CTC CAG GCT AAA ATT ACA        335
Met Gln Gln Gln Asp Ser Ala Ile Gln Asn Leu Gln Ala Lys Ile Thr
                100                 105                 110

GAC TTG GAA TCA CAA GTA AGT GAA GCC GTT AGA TCT GAC ACA ACA AGA        383
Asp Leu Glu Ser Gln Val Ser Glu Ala Val Arg Ser Asp Thr Thr Arg
            115                 120                 125

ACA GGA GAT GCC TTG CAA TCT CAG GAC ATA TTT TCT CCA ATA CCA AAA        431
Thr Gly Asp Ala Leu Gln Ser Gln Asp Ile Phe Ser Pro Ile Pro Lys
        130                 135                 140

GCG GTT GAG GGT ACA ACT GAT TCT TCT TCT GTT ACC AAG AAA CTT GAG        479
Ala Val Glu Gly Thr Thr Asp Ser Ser Ser Val Thr Lys Lys Leu Glu
    145                 150                 155

GAA GAA TTG AAA AAA CGT GAT GCA CTG ATT GAG AGG TTG CAT GAA GAA        527
Glu Glu Leu Lys Lys Arg Asp Ala Leu Ile Glu Arg Leu His Glu Glu
160                 165                 170                 175

AAC GAA AAG TTG TTT GAC AGA TTA ACA GAA AGG TCA ATG GCT GTT TCG        575
Asn Glu Lys Leu Phe Asp Arg Leu Thr Glu Arg Ser Met Ala Val Ser
                180                 185                 190

ACC CAG GTG TTG AGT CCC TCA TTA AGA GCT TCG CCA AAC ATT CAG CCT        623
Thr Gln Val Leu Ser Pro Ser Leu Arg Ala Ser Pro Asn Ile Gln Pro
            195                 200                 205

GCC AAT GTT AAC AGG GGT GAA GGA TAT TCG GCA GAA GCC GTT GCT TTA        671
Ala Asn Val Asn Arg Gly Glu Gly Tyr Ser Ala Glu Ala Val Ala Leu
        210                 215                 220

CCA TCT ACA CCA AAT AAG AAT AAC GGA GCG ATT ACG TTA GTA AAA TCT        719
Pro Ser Thr Pro Asn Lys Asn Asn Gly Ala Ile Thr Leu Val Lys Ser
    225                 230                 235

GGC ACT GAT TTA GTA AAA ACC ACT CCA GCT GGA GAA TAC TTG ACA GCT        767
Gly Thr Asp Leu Val Lys Thr Thr Pro Ala Gly Glu Tyr Leu Thr Ala
240                 245                 250                 255
```

```
GCA TTG AAT GAC TTT GAC CCT GAA GAA TAT GAA GGT CTT GCT GCC ATT        815
Ala Leu Asn Asp Phe Asp Pro Glu Glu Tyr Glu Gly Leu Ala Ala Ile
                260                 265                 270

GCT GAC GGC GCA AAC AAG CTA CTA ATG CTG GTT TTG GCA GCT GTC ATC        863
Ala Asp Gly Ala Asn Lys Leu Leu Met Leu Val Leu Ala Ala Val Ile
            275                 280                 285

AAG GCT GGT GCT TCC AGA GAG CAT GAA ATC CTT GCT GAG ATT AGA GAT        911
Lys Ala Gly Ala Ser Arg Glu His Glu Ile Leu Ala Glu Ile Arg Asp
        290                 295                 300

TCT GTC TTT TCA TTT ATT CGG AAA ATG GAA CCA AGA AGA GTA ATG GAT        959
Ser Val Phe Ser Phe Ile Arg Lys Met Glu Pro Arg Arg Val Met Asp
    305                 310                 315

ACC ATG CTT GTT TCC CGA GTT AGG ATA CTA TAC ATA AGG TCC TTA CTG       1007
Thr Met Leu Val Ser Arg Val Arg Ile Leu Tyr Ile Arg Ser Leu Leu
320                 325                 330                 335

GCA CGA TCA CCG GAG CTT CAG ACT ATC AGG GTC TCT CCT GTC GAG TGC       1055
Ala Arg Ser Pro Glu Leu Gln Thr Ile Arg Val Ser Pro Val Glu Cys
                340                 345                 350

TTT CTT GAG AAG CCT AAT ACT GGT AGA AGT AAA AGC ACT AGC AGG GGT       1103
Phe Leu Glu Lys Pro Asn Thr Gly Arg Ser Lys Ser Thr Ser Arg Gly
            355                 360                 365

AGC AGC CCA GGT AGA TCC CCT GTT CGA TAT CTT GAT ACG CAG ATC CAT       1151
Ser Ser Pro Gly Arg Ser Pro Val Arg Tyr Leu Asp Thr Gln Ile His
        370                 375                 380

GGC TTT AAA GTA AAT ATA AAG GCA GAA AGG AGG                          1184
Gly Phe Lys Val Asn Ile Lys Ala Glu Arg Arg
    385                 390
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 394 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Ala Ser Asp Ala Arg Lys Glu Leu Leu Glu Lys Glu Arg Glu Asn Gln
 1               5                  10                  15

Asn Leu Lys Gln Glu Val Val Gly Leu Lys Lys Ala Leu Lys Asp Ala
            20                  25                  30
```

Asn Asp Gln Cys Val Leu Leu Tyr Ser Glu Val Gln Arg Ala Trp Lys
35 40 45

Val Ser Phe Thr Leu Gln Ser Asp Leu Lys Ser Glu Asn Ile Met Leu
50 55 60

Val Asp Lys His Arg Leu Glu Lys Glu Gln Asn Ser Gln Leu Arg Asn
65 70 75 80

Gln Ile Ala Gln Phe Leu Gln Leu Asp Gln Glu Gln Lys Leu Gln Met
85 90 95

Gln Gln Gln Asp Ser Ala Ile Gln Asn Leu Gln Ala Lys Ile Thr Asp
100 105 110

Leu Glu Ser Gln Val Ser Glu Ala Val Arg Ser Asp Thr Thr Arg Thr
115 120 125

Gly Asp Ala Leu Gln Ser Gln Asp Ile Phe Ser Pro Ile Pro Lys Ala
130 135 140

Val Glu Gly Thr Thr Asp Ser Ser Ser Val Thr Lys Lys Leu Glu Glu
145 150 155 160

Glu Leu Lys Lys Arg Asp Ala Leu Ile Glu Arg Leu His Glu Glu Asn
165 170 175

Glu Lys Leu Phe Asp Arg Leu Thr Glu Arg Ser Met Ala Val Ser Thr
180 185 190

Gln Val Leu Ser Pro Ser Leu Arg Ala Ser Pro Asn Ile Gln Pro Ala
195 200 205

Asn Val Asn Arg Gly Glu Gly Tyr Ser Ala Glu Ala Val Ala Leu Pro
210 215 220

Ser Thr Pro Asn Lys Asn Asn Gly Ala Ile Thr Leu Val Lys Ser Gly
225 230 235 240

Thr Asp Leu Val Lys Thr Thr Pro Ala Gly Glu Tyr Leu Thr Ala Ala
245 250 255

Leu Asn Asp Phe Asp Pro Glu Glu Tyr Glu Gly Leu Ala Ala Ile Ala
260 265 270

Asp Gly Ala Asn Lys Leu Leu Met Leu Val Leu Ala Ala Val Ile Lys
275 280 285

Ala Gly Ala Ser Arg Glu His Glu Ile Leu Ala Glu Ile Arg Asp Ser
290 295 300

Val Phe Ser Phe Ile Arg Lys Met Glu Pro Arg Arg Val Met Asp Thr
305 310 315 320

Met Leu Val Ser Arg Val Arg Ile Leu Tyr Ile Arg Ser Leu Leu Ala
325 330 335

Arg Ser Pro Glu Leu Gln Thr Ile Arg Val Ser Pro Val Glu Cys Phe
340 345 350

Leu Glu Lys Pro Asn Thr Gly Arg Ser Lys Ser Thr Ser Arg Gly Ser
355 360 365

Ser Pro Gly Arg Ser Pro Val Arg Tyr Leu Asp Thr Gln Ile His Gly
370 375 380

Phe Lys Val Asn Ile Lys Ala Glu Arg Arg
385 390

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CGAGATCTGA ATTCATGGAT CAGTA 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGAGATCTGA ATTCCTAAGG CATGCC 26

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CGGATCCGAA TTCATGGAGA ACGAG 25

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGGATCCGAA TTCTCAGAAC TGAGA 25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ATACCACTAC AATGGATG 18

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AGTTGAAGTG AACTTGCGGG 20